



OIPE

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/258,600

DATE: 10/28/2002

TIME: 15:04:48

Input Set : N:\Crf3\RULE60\09258600.raw
 Output Set: N:\CRF4\10282002\I258600.raw

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: FOWLKES, Dana M.
 6 BROACH, Jim
 7 MANFREDI, John
 8 KLEIN, Christine
 9 MURPHY, Andrew J.
 10 PAUL, Jeremy
 11 TRUEHEART, Joshua

ENTERED

13 (ii) TITLE OF INVENTION: YEAST CELLS ENGINEERED TO PRODUCE
 14 PHEROMONE SYSTEM PROTEIN SURROGATES, AND USES THEREFOR

16 (iii) NUMBER OF SEQUENCES: 119

18 (iv) CORRESPONDENCE ADDRESS:

19 (A) ADDRESSEE: BROWDY AND NEIMARK
 20 (B) STREET: 419 Seventh Street, N.W., Suite 300
 21 (C) CITY: Washington
 22 (D) STATE: D.C.
 23 (E) COUNTRY: USA
 24 (F) ZIP: 20004

26 (v) COMPUTER READABLE FORM:

27 (A) MEDIUM TYPE: Floppy disk
 28 (B) COMPUTER: IBM PC compatible
 29 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 30 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

32 (vi) CURRENT APPLICATION DATA:

33 (A) APPLICATION NUMBER: US/09/258,600

C--> 34 (B) FILING DATE: 26-Feb-1999

35 (C) CLASSIFICATION:

53 (vii) PRIOR APPLICATION DATA:

38 (A) APPLICATION NUMBER: US/08/461,598

39 (B) FILING DATE: 05-JUN-1995

42 (A) APPLICATION NUMBER: US 08/322,137

43 (B) FILING DATE: 13-OCT-1994

46 (A) APPLICATION NUMBER: US 08/309,313

47 (B) FILING DATE: 20-SEP-1994

50 (A) APPLICATION NUMBER: US 08/190,328

51 (B) FILING DATE: 31-JAN-1994

54 (A) APPLICATION NUMBER: US 08/041,431

55 (B) FILING DATE: 31-MAR-1993

57 (viii) ATTORNEY/AGENT INFORMATION:

58 (A) NAME: COOPER, Iver P.

59 (B) REGISTRATION NUMBER: 28,005

60 (C) REFERENCE/DOCKET NUMBER: FOLWKES=2F

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/258,600

DATE: 10/28/2002

TIME: 15:04:48

Input Set : N:\Crf3\RULE60\09258600.raw
 Output Set: N:\CRF4\10282002\I258600.raw

62 (ix) TELECOMMUNICATION INFORMATION:

 63 (A) TELEPHONE: 202-628-5197
 64 (B) TELEFAX: 202-737-3528
 65 (C) TELEX: 248633

68 (2) INFORMATION FOR SEQ ID NO: 1:

 70 (i) SEQUENCE CHARACTERISTICS:
 71 (A) LENGTH: 89 amino acids
 72 (B) TYPE: amino acid
 73 (C) STRANDEDNESS: single
 74 (D) TOPOLOGY: linear

76 (ii) MOLECULE TYPE: peptide

79 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

 81 Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser
 82 1 5 10 15
 84 Ala Leu Ala Ala Pro Val Asn Thr Thr Glu Asp Glu Thr Ala Gln
 85 20 25 30
 87 Ile Pro Ala Glu Ala Val Ile Gly Tyr Leu Asp Leu Glu Gly Asp Phe
 88 35 40 45
 90 Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu
 91 50 55 60
 93 Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val
 94 65 70 75 80
 96 Ser Leu Asp Lys Arg Glu Ala Glu Ala
 97 85

99 (2) INFORMATION FOR SEQ ID NO: 2:

 101 (i) SEQUENCE CHARACTERISTICS:
 102 (A) LENGTH: 76 amino acids
 103 (B) TYPE: amino acid
 104 (D) TOPOLOGY: linear

106 (ii) MOLECULE TYPE: peptide

109 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

 111 Trp His Trp Leu Gln Leu Lys Pro Gly Gln Pro Met Tyr Lys Arg Glu
 112 1 5 10 15
 114 Ala Glu Ala Glu Ala Trp His Trp Leu Gln Leu Lys Pro Gly Gln Pro
 115 20 25 30
 117 Met Tyr Lys Arg Glu Ala Asp Ala Glu Ala Trp His Trp Leu Gln Leu
 118 35 40 45
 120 Lys Pro Gly Gln Pro Met Tyr Lys Arg Glu Ala Asp Ala Glu Ala Trp
 121 50 55 60
 123 His Trp Leu Gln Leu Lys Pro Gly Gln Pro Met Tyr
 124 65 70 75

126 (2) INFORMATION FOR SEQ ID NO: 3:

 128 (i) SEQUENCE CHARACTERISTICS:
 129 (A) LENGTH: 15 base pairs
 130 (B) TYPE: nucleic acid
 131 (C) STRANDEDNESS: double
 132 (D) TOPOLOGY: linear

W--> 134 (ii) MOLECULE TYPE: synthetic DNA

137 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/258,600

DATE: 10/28/2002

TIME: 15:04:48

Input Set : N:\Crf3\RULE60\09258600.raw
 Output Set: N:\CRF4\10282002\I258600.raw

139 AAGCTTAAAAA GAATG 15
 141 (2) INFORMATION FOR SEQ ID NO: 4:
 143 (i) SEQUENCE CHARACTERISTICS:
 144 (A) LENGTH: 37 base pairs
 145 (B) TYPE: nucleic acid
 146 (C) STRANDEDNESS: single
 147 (D) TOPOLOGY: linear
 149 (ii) MOLECULE TYPE: cDNA
 152 (ix) FEATURE:
 153 (A) NAME/KEY: CDS
 154 (B) LOCATION: 1..24
 157 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:
 159 AAA GAA GAA GGG GIA TCT TTG CTT AAGCTCGAGA TCT 37
 160 Lys Glu Glu Gly Val Ser Leu Leu
 161 1 5
 164 (2) INFORMATION FOR SEQ ID NO: 5:
 166 (i) SEQUENCE CHARACTERISTICS:
 167 (A) LENGTH: 8 amino acids
 168 (B) TYPE: amino acid
 169 (D) TOPOLOGY: linear
 171 (ii) MOLECULE TYPE: peptide
 173 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:
 175 Lys Glu Glu Gly Val Ser Leu Leu
 176 1 5
 178 (2) INFORMATION FOR SEQ ID NO: 6:
 180 (i) SEQUENCE CHARACTERISTICS:
 181 (A) LENGTH: 77 base pairs
 182 (B) TYPE: nucleic acid
 183 (C) STRANDEDNESS: double
 184 (D) TOPOLOGY: linear
 W--> 186 (ii) MOLECULE TYPE: synthetic DNA
 189 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
 191 CGTGAAGCTT AAGCGTGAGG CAGAAGCTNN KNNKNNKNNK NKNKNNKNNK NKNKNNKNN 60
 193 KNNKNNKNTGA TCATCCG 77
 195 (2) INFORMATION FOR SEQ ID NO: 7:
 197 (i) SEQUENCE CHARACTERISTICS:
 198 (A) LENGTH: 19 amino acids
 199 (B) TYPE: amino acid
 200 (D) TOPOLOGY: linear
 202 (ii) MOLECULE TYPE: peptide
 205 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
 W--> 207 Lys Arg Glu Ala Glu Ala Xaa 15
 208 1 5 10
 W--> 210 Xaa Xaa Xaa
 214 (2) INFORMATION FOR SEQ ID NO: 8:
 216 (i) SEQUENCE CHARACTERISTICS:
 217 (A) LENGTH: 36 amino acids
 218 (B) TYPE: amino acid
 219 (D) TOPOLOGY: linear

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/258,600

DATE: 10/28/2002

TIME: 15:04:48

Input Set : N:\Crf3\RULE60\09258600.raw
 Output Set: N:\CRF4\10282002\I258600.raw

221 (ii) MOLECULE TYPE: peptide
 224 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
 226 Met Gln Pro Ser Thr Ala Thr Ala Ala Pro Lys Glu Lys Thr Ser Ser
 227 1 5 10 15
 229 Glu Lys Lys Asp Asn Tyr Ile Ile Lys Gly Val Phe Trp Asp Pro Ala
 230 20 25 30
 232 Cys Val Ile Ala
 233 35

235 (2) INFORMATION FOR SEQ ID NO: 9:
 237 (i) SEQUENCE CHARACTERISTICS:
 238 (A) LENGTH: 19 base pairs
 239 (B) TYPE: nucleic acid
 240 (C) STRANDEDNESS: single
 241 (D) TOPOLOGY: linear

W--> 243 (ii) MOLECULE TYPE: synthetic DNA
 246 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

19

248 AAGCTTCCGA ATAGAAATG
 250 (2) INFORMATION FOR SEQ ID NO: 10:
 252 (i) SEQUENCE CHARACTERISTICS:
 253 (A) LENGTH: 36 base pairs
 254 (B) TYPE: nucleic acid
 255 (C) STRANDEDNESS: double
 256 (D) TOPOLOGY: linear

W--> 258 (ii) MOLECULE TYPE: synthetic DNA
 261 (ix) FEATURE:

262 (A) NAME/KEY: CDS
 263 (B) LOCATION: 1..27
 266 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

36

268 GCC GCT CCA AAA GAA AAG ACC TCG AGC TCGCTTAAG
 269 Ala Ala Pro Lys Glu Lys Thr Ser Ser
 270 1 5

273 (2) INFORMATION FOR SEQ ID NO: 11:
 275 (i) SEQUENCE CHARACTERISTICS:
 276 (A) LENGTH: 9 amino acids
 277 (B) TYPE: amino acid
 278 (D) TOPOLOGY: linear

280 (ii) MOLECULE TYPE: peptide
 282 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

284 Ala Ala Pro Lys Glu Lys Thr Ser Ser
 285 1 5

287 (2) INFORMATION FOR SEQ ID NO: 12:
 289 (i) SEQUENCE CHARACTERISTICS:
 290 (A) LENGTH: 79 base pairs
 291 (B) TYPE: nucleic acid
 292 (C) STRANDEDNESS: double
 293 (D) TOPOLOGY: linear

295 (ii) MOLECULE TYPE: cDNA
 298 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

300 GGTACTCGAG TGAAAAGAAG GACAACNNKN NKNNKNNKNN KNNKNNKNNK NNKNNKNNKT

60

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/258,600

DATE: 10/28/2002
TIME: 15:04 48

Input Set : N:\CrF3\RULE60\09258600.raw
Output Set: N:\CRF4\10282002\I258600.raw

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/258,600

DATE: 10/28/2002
TIME: 15:04:49

Input Set : N:\Crf3\RULE60\09258600.raw
Output Set: N:\CRF4\10282002\I258600.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:6; N Pos. 29,30,32,33,35,36,38,39,41,42,44,45,47,48,50,51,53,54,56,57

Seq#:6; N Pos. 59,60,62,63,65,66

Seq#:7; Xaa Pos.7,8,9,10,11,12,13,14,15,16,17,18,19

Seq#:12; N Pos. 27,28,30,31,33,34,36,37,39,40,42,43,45,46,48,49,51,52,54,55

Seq#:12; N Pos. 57,58

Seq#:13; Xaa Pos.8,9,10,11,12,13,14,15,16,17,18

Seq#:27; N Pos. 12,13,15,16,18,19,21,22,24,25,27,28,30,31,33,34,36,37,39,40

Seq#:27; N Pos. 42,43,45,46,48,49

Seq#:29; N Pos. 22,23,25,26,28,29,31,32,34,35,37,38,40,41,43,44,46,47,49,50

Seq#:29; N Pos. 52,53

Seq#:39; N Pos. 19,20,22,23,25,26,28,29,31,32,34,35,37,38,40,41,43,44,46,47

Seq#:39; N Pos. 49,50,52,53

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/258,600

DATE: 10/28/2002

TIME: 15:04:49

Input Set : N:\Crf3\RULE60\09258600.raw
Output Set: N:\CRF4\10282002\I258600.raw

L:33 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:34 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:134 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=3
L:186 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=6
L:207 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:0
L:210 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:16
L:243 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=9
L:258 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=10
L:316 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:0
L:319 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13 after pos.:17
L:331 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=14
L:346 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=15
L:361 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=16
L:376 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=17
L:391 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=18
L:407 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=19
L:422 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=20
L:437 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=21
L:456 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=22
L:471 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=23
L:501 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=25
L:520 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=26
L:550 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=28
L:565 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=29
L:581 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=30
L:596 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=31
L:611 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=32
L:626 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=33
L:661 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=35
L:676 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=36
L:852 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:46
L:889 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:48
L:926 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:50
L:963 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:52
L:1000 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:54
L:1037 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:56
L:1681 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:91
L:1718 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:93
L:1755 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:95
L:1792 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:97
L:1829 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:99
L:1866 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:101
L:1903 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:103
L:1940 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:105
L:1977 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:107
L:2014 M:336 W: Invalid Amino Acid Number in Coding Region, SEQ ID:109
L:2217 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=118
L:2236 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=119